UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/684,508	10/15/2003	Livia Polanyi	CQ10224	6736

23493 7590 02/05/2008

SUGHRUE MION, PLLC
2100 Pennsylvania Avenue, N.W.
Washington, DC 20037

EXAMINER
COLUCCI, MICHAEL C

ART UNIT	PAPER NUMBER
2626	

NOTIFICATION DATE	DELIVERY MODE
02/05/2008	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

USPTO@sughrue.com
USPatDocketing@sughrue.com

Office Action Summary	Application No.	Applicant(s)	
	10/684,508	POLANYI ET AL.	
	Examiner	Art Unit	
	Michael C. Colucci	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *14 November 2007*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-27,29-32 and 34-39* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☒ Claim(s) *8,18,34,35 and 37* is/are allowed.

6)☒ Claim(s) *1-4, 9-14,19-27,29-32,36,38 and 39* is/are rejected.

7)☒ Claim(s) *5-7 and 15-17* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *15 October 2003* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *See Continuation Sheet*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

Continuation of Attachment(s) 3). Information Disclosure Statement(s) (PTO/SB/08), Paper No(s)/Mail Date :03/20/2006, 08/02/2004, 10/15/2003.

DETAILED ACTION

Allowable Subject Matter

1. Claims 8, 18, 34, 35, and 37 allowed.

2. Claims 5-7 and 15-17 objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

Response to Arguments

3. Applicant's arguments, see (Remarks pages 21-22), filed 11/14/2007, with respect to claims 5-8, 15-18, 34-35, and 37 have been fully considered and are persuasive. The rejection of claims 5-8, 15-18, 34-35, and 37 has been withdrawn.

4. Applicant's arguments filed 11/14/2007 (Remarks pages 18-19), with respect to the rejection(s) of claim(s) 24, 29, and 38 under 102(b) have been fully considered but they are not persuasive.

Re claims 24, 29, and 38, mapping between sentential level parse features and discourse level parse feature is construed to be equally effective and functionally equivalent as tagging features within text such as commas, parenthesis, dashes, etc. as well as parts of speech tagging. The concept of mapping is not found within the specification as to further narrow the limitation and is therefore broadly interpreted to have the functionality of tagging. Additionally, discourse level data s construed to be functionally equivalent and equally effective to sentential level data, where syntactic

relations are a form of mapping, where data within text can be mapped to other data within text as to abide by semantic and/or syntactic rules.

Claim Rejections - 35 USC § 102

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

6. Claims 24-27, 29-32, and 38-39 rejected under 35 U.S.C. 102(b) as being anticipated by Marcu et al U.S. PGPUB 20020046018 (herein after Marcu).

Re claims 24, 29, and 38, Marcu teaches a method for discourse parsing comprising the steps of:

determining a structural representation of discourse ([0003]) based on a theory of discourse analysis ([0077]);

determining at least one sentence of a text ([0010]);

determining sentential-level parse features for the at least one sentence ([0100]);

determining a mapping between the sentential-level parse features and discourse-level parse features ([0100]);

determining a discourse-level parse tree of the at least one sentence based on the mapping ([0119]);

determining a main discourse Constituent for the at least one sentence ([0100]);

determining an attachment of the discourse level parse tree ([0119]) to the structural representation of discourse by the determined main discourse constituent based on attachment rules ([0098]) for the theory of discourse ([0077]).

Re claims 25, 30, and 39 Marcu teaches a method of segmenting text into discourse constituents comprising the steps of:

determining a theory of discourse analysis ([0077]);

determining candidate segments ([0223]);

determining attributes of candidate segments associated with continuing the discourse ([0100]);

determining if the candidate segment is a discourse constituent ([0119]) based on the theory of discourse analysis ([0077]) and the determined attributes ([0100]).

Re claims 26 and 31, Marcu teaches the method of claim 25, wherein the attributes are determined based on at least one of: a part-of-speech tag, a probabilistic parser, a statistical parser, a finite state parser, a symbolic parser, a lexicon, and a WordNet relation ([0100]).

Re claims 27 and 32, Marcu teaches a method of determining a structural representation of discourse comprising the steps of:

determine discourse constituents for a text ([0077]);

conjoining the discourse constituents into a structural representation of discourse ([0119]) based on theory of discourse analysis classifications ([0077]) of the discourse constituents and at least one of a syntactic, a semantic and a lexical-semantic constraint ([0021]).

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. Claims 1-4, 11-14, 22-23, 36 rejected under 35 U.S.C. 103(a) as being unpatentable over Marcu et al U.S. PGPUB 20020046018 (herein after Marcu) in view of Copperman et al US PGPUB 20040024739 A1 (hereinafter Copperman).

Re claims 1, 11, 22-23, and 36, Marcu teaches a method of determining a hybrid text summary comprising the steps

determining discourse constituents for a text ([0010]);

determining a structural representation of discourse for the text ([0003]);

discourse constituents based on at least one non-structural measure of relevance ([0003]);

determining a hybrid text summary ([0003])

However, Marcu fails to teach determining relevance scores (Copperman [0176])

percolating (Copperman [0165] & Fig. 18) relevance scores (Copperman [0176]) based on the structural representation of discourse (Copperman [0082]);

based on discourse constituents with relevance scores compared to a threshold relevance score (Copperman [0082]).

NOTE: For purposes of prior art, percolation is construed to be both functionally equivalent and equally effective as passing various scores with probability, where undesirable scores can be filtered out.

Copperman teaches search that is performed by invoking a content-based search engine one or more times, each time specifying a query and some set of indexes. Conceptually, the search engine is applied separately for each region. Regions are formed dynamically, and the objects on which search engines function are statically built indexes. Therefore, calling the search engine on a region is realized in approximation: for each region, a covering set of indexes is found from the mapping of nodes to indexes. More specifically, as shown in FIG. 18, taxonomy 1500 comprises regions 1510 and 1520. Region 1510 is comprised entirely of the green cluster (FIG. 14) so the search on this region would be limited to index 1150. Region 1520, on the other hand, comprises the orange cluster (FIG. 14) and the purple cluster (FIG. 14). Therefore, a search on this region would have to include indexes 1145 and 1155.

Additionally, Copperman teaches threshold scores for tagging either document-knowledge containers or question-knowledge containers, and maximum numbers of tags to assign from each topic taxonomy to either document-knowledge containers or

question-knowledge containers. Copperman teaches a list of result knowledge containers that are possible "answers" to the question, each with a relevance score between 0 and 1.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention determining a hybrid text based on constituents with relevance scores having a threshold score, where percolation is performed based on a relevance score. Percolating or passing relevance scores relevant to a threshold score would allow for efficient clustering, where only a specific range of data that abides by a relevance probability will remain.

Re claims 2 and 12, Marcu teaches the method of claim 1, wherein the structural representation of discourse ([0003]) is determined based on a theory of discourse analysis ([0077]).

Re claims 3 and 13, Marcu teaches the method of claim 2, wherein the theory of discourse analysis is at least one of the Linguistic Discourse Model, the Unified Linguistic Discourse Model, Rhetorical Structure Theory, Discourse Structure Theory and Structured Discourse Representation Theory ([0077]).

Re claims 4 and 14, Marcu teaches the method of claim 1, wherein non-structural measures of relevance are determined based on at least one of statistics ([0030]), keywords ([0023]), knowledge bases ([0067]).

9. Claims 9-10 and 19-21 rejected under 35 U.S.C. 103(a) as being unpatentable over Marcu et al U.S. PGPUB 20020046018 (herein after Marcu) in view of Copperman et al US PGPUB 20040024739 A1 (hereinafter Copperman) and further in view of Nakao US 6205456 B1 (hereinafter Nakao).

Re claims 9 and 19-20, Marcu fails to teach the method of claim 1, further comprising the steps of:

discourse constituents with relevance scores more relevant that the threshold relevance score (Copperman [0082]).

Copperman teaches search that is performed by invoking a content-based search engine one or more times, each time specifying a query and some set of indexes. Conceptually, the search engine is applied separately for each region. Regions are formed dynamically, and the objects on which search engines function are statically built indexes. Therefore, calling the search engine on a region is realized in approximation: for each region, a covering set of indexes is found from the mapping of nodes to indexes. More specifically, as shown in FIG. 18, taxonomy 1500 comprises regions 1510 and 1520. Region 1510 is comprised entirely of the green cluster (FIG. 14) so the search on this region would be limited to index 1150. Region 1520, on the other hand, comprises the orange cluster (FIG. 14) and the purple cluster (FIG. 14). Therefore, a search on this region would have to include indexes 1145 and 1155.

Copperman also teaches allowing users to select among regions, where the system may allow users to adjust regions. This can involve either adding or removing

concept-nodes to/from a region that has been identified for the question. For example, suppose the system believes a user's question is about sports and during one step of the dialog returns a taxonomic region containing a general "Sports" concept-node and a variety of descendent concept-nodes for different types of sports.

Additionally, Copperman teaches threshold scores for tagging either document-knowledge containers or question-knowledge containers, and maximum numbers of tags to assign from each topic taxonomy to either document-knowledge containers or question-knowledge containers. Copperman teaches a list of result knowledge containers that are possible "answers" to the question, each with a relevance score between 0 and 1.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention determining a hybrid text based on constituents with relevance scores having a threshold score, where threshold comparison is performed based on a relevance score. Comparing relevance scores relevant to a threshold score would allow for efficient clustering, where only a specific range of data that abides by a relevance probability will remain.

However, Marcu in view of Copperman fails to teach determining every leaf discourse constituent containing an anaphor (Nakao col 24 line 46-57);

for each anaphor, determine any unique antecedent referents for the anaphor (Nakao col 4 line 12-29);

substituting the unique antecedent referent into the leaf discourse constituent for the anaphor (Nakao col 4 line 12-29);

removing the discourse constituent (Nakao col 25 lines 33-50) containing the unique antecedent referent from the set (Nakao col 4 line 12-29).

Nakao teaches for an anaphoric expression, its antecedent is searched for and the anaphoric expression is replaced with the antecedent or a portion containing the antecedent is included in a summary so that the summary can be easily understood. The antecedent of the anaphoric expression can be identified by a method referred to as a centering method. This method makes a list of centers that comprises probable elements (centers) of a sentence to be antecedents of anaphoric expressions in the subsequent sentences. The elements probability to be an antecedent is calculated mainly by its syntactic role in a sentence, such as subject, direct object, etc. Then, the method resolves an anaphoric expression by selecting the most probable element from the list with the restriction of agreement of number, gender, etc.

NOTE: For purposes of prior art, a reduced span is construed to be both functionally equivalent and equally effective as a resolved set with the most probable elements present.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention determining portions of text containing anaphors, where an antecedent referent can be substituted for the anaphor, and removing the antecedent referent relevant to a relevancy threshold. Using an anaphor to replace an antecedent

referent would allow for a preservation of memory, where more than one item can refer to a single item (i.e. anaphoric relationship). Additionally, using syntactic analysis with antecedent and anaphoric substitutions would allow for a minimized set of data that would the most probable with respect to the discourse/text by reduction of a set of candidate data. Additionally, scoring various elements relevant to probability and antecedent basis would allow for a further reduced set that is minimized for both anaphor and original data from the discourse.

Re claims 10 and 21, method of claim 1, wherein percolation of relevance scores comprises the steps of:

percolating relevance score of important discourse constituents (Copperman [0165] & Fig. 18).

Copperman teaches search that is performed by invoking a content-based search engine one or more times, each time specifying a query and some set of indexes. Conceptually, the search engine is applied separately for each region. Regions are formed dynamically, and the objects on which search engines function are statically built indexes. Therefore, calling the search engine on a region is realized in approximation: for each region, a covering set of indexes is found from the mapping of nodes to indexes. More specifically, as shown in FIG. 18, taxonomy 1500 comprises regions 1510 and 1520. Region 1510 is comprised entirely of the green cluster (FIG. 14) so the search on this region would be limited to index 1150. Region 1520, on the

other hand, comprises the orange cluster (FIG. 14) and the purple cluster (FIG. 14). Therefore, a search on this region would have to include indexes 1145 and 1155.

Copperman also teaches allowing users to select among regions, where the system may allow users to adjust regions. This can involve either adding or removing concept-nodes to/from a region that has been identified for the question. For example, suppose the system believes a user's question is about sports and during one step of the dialog returns a taxonomic region containing a general "Sports" concept-node and a variety of descendent concept-nodes for different types of sports.

Additionally, Copperman teaches threshold scores for tagging either document-knowledge containers or question-knowledge containers, and maximum numbers of tags to assign from each topic taxonomy to either document-knowledge containers or question-knowledge containers. Copperman teaches a list of result knowledge containers that are possible "answers" to the question, each with a relevance score between 0 and 1.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention determining a hybrid text based on constituents with percolating relevance scores having a threshold score, where threshold comparison is performed based on a relevance score. Percolating relevance scores relevant to a threshold score would allow for efficient clustering, where only a specific range of data that abides by a relevance probability will remain.

determining important discourse constituent nodes; determining unresolved anaphors (Nakao col 4 line 12-29);

determining potential resolving discourse constituents containing potential antecedent referent with potential to resolve anaphors (Nakao col 4 line 12-29);

through a reduced span of potential resolving discourse constituents (Nakao col 4 line 12-29);

determining a reduced span of discourse constituents based on relevance score (Nakao col 4 line 12-29).

Nakao teaches for an anaphoric expression, its antecedent is searched for and the anaphoric expression is replaced with the antecedent or a portion containing the antecedent is included in a summary so that the summary can be easily understood. The antecedent of the anaphoric expression can be identified by a method referred to as a centering method. This method makes a list of centers that comprises probable elements (centers) of a sentence to be antecedents of anaphoric expressions in the subsequent sentences. The elements probability to be an antecedent is calculated mainly by its syntactic role in a sentence, such as subject, direct object, etc. Then, the method resolves an anaphoric expression by selecting the most probable element from the list with the restriction of agreement of number, gender, etc.

NOTE: For purposes of prior art, a reduced span is construed to be both functionally equivalent and equally effective as a resolved set with the most probable elements present.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention determining portions of text containing anaphors, where an antecedent referent can be substituted for the anaphor, and removing the antecedent referent relevant to a relevancy threshold. Using an anaphor to replace an antecedent referent would allow for a preservation of memory, where more than one item can refer to a single item (i.e. anaphoric relationship). Additionally, using syntactic analysis with antecedent and anaphoric substitutions would allow for a minimized set of data that would the most probable with respect to the discourse/text by reduction of a set of candidate data. Additionally, percolating scores of various elements relevant to probability and antecedent basis would allow for a further reduced set that is minimized for both anaphor and original data from the discourse.

Conclusion

10. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. US 20040153440 A1, US 20040139057 A1, US 20040034520 A1, US 20030093514 A1, US 20020052730 A1, US 20040054654 A1, US 20020040363 A1, US 6349316 B1, US 20040044519 A1, US 6571238 B1, US 6188976 B1, US 6738759 B1, US 20070239660 A1, US 20040030741 A1, US 20020198697 A1.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-

270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Michael Colucci Jr.
Patent Examiner
AU 2626
(571)-270-1847
Michael.Colucci@uspto.gov

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER